National Biological Information Infrastructure

www.nbii.gov

The NBII Biological Metadata Standard

A key element in fostering development of the National Biological Information Infrastructure (NBII) as a distributed federation of biological data and information is the availability of a standardized format to use in describing these data and information sources (including subject matter; how, when, where, and by whom they were collected; accuracy; and quality) so people can quickly and easily compare and contrast among many different sources to choose those that best meet their needs. A standard "metadata" format works in the same way that the uniformly presented information in a library's card catalog helps you find a particular book or magazine article — quickly! These attributes help explain why this biological metadata standard is a fundamental component of the overall NBII effort.

The U.S. Geological Survey's Biological Resources Division (BRD) has developed a biological "profile" of the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata that increases the utility of this standard for documenting biological resources data and information. This biological metadata standard will support increased access to, sharing, and use of biological data and information among users on both a national and an international basis. It will also help to broaden the understanding and implementation of the FGDC geospatial metadata content standard within the biological resources community.

Although the FGDC metadata content standard provides excellent documentation of a data set from the geospatial perspective, it is limited and, in some aspects, inadequate, for describing data from the biological perspective. For example, one consistent need in describing different biological data sets so that they can be compared, shared, and integrated is to document the necessary bio-systematics aspects of the data (that is, taxonomy and nomenclature for species and higher taxonomic groups). Obviously, the existing FGDC geospatial metadata standard was not designed to focus on this type of community-specific requirement.

In addition, a significant portion of the available body of biological data and information was not collected or intended to be applied in a way that is explicitly geospatial; for example, data resulting from in vitro laboratory research or directories of biological science experts. Rather than document these types of "non-geospatial" data and information using an entirely different metadata content standard, BRD wanted to provide—as part of the NBII—an inclusive biological metadata standard that could be used to describe biological data and information in all its forms, and also provide maximum compatibility with the FGDC geospatial standard.

In developing the biological metadata content standard, the BRD actively sought the involvement of the biological science community, both inside and outside of government. This included soliciting the participation of the American Institute of Biological Sciences, which convened a national panel of biologists to review and comment on the work of the BRD in

NBII National Program Office
U.S. Geological Survey, 302 National Center
Reston, Virginia 20192

Phone: 703-648-4205
Fax: 703-648-4224
E-mail: nbii_program@nbii.gov

developing a draft biological metadata standard. The resulting standard (http://www.nbii.gov/current.status.html) functions as a specialized disciplinary "profile" or superset of the FGDC geospatial metadata standard, including all its elements while adding key biological science elements in three areas:

1) Descriptions of the field and/or laboratory methodologies used to collect the data,

2) Descriptions of any analytical tools that have been applied to, or are needed to interpret the data, and

3) Taxonomic reference information (for example, taxonomic coverage/extent, sources and references for species nomenclature, and treatment of voucher specimens).

All three areas have obvious importance in terms of allowing the holders of biological data and information to adequately describe their material, and for the potential users of these data and information to make informed decisions as to whether a given source satisfies their requirements.

The NBII biological metadata standard can be used by anyone interested in increasing access to biological data or information. This includes explicitly geospatial biological data; biological data that are not explicitly geospatial, such as those resulting from in vitro research; and biological information, such as research publications, bibliographies, lists of experts, and specimen collections. The biological metadata standard is hardware and software independent. Any metadata created in compliance with this profile will function in harmony with metadata created with the "base" FGDC geospatial metadata standard —and with associated metadata search, indexing, and exchange tools, formats, and protocols.

The impact of the development of this standard biological "profile" of the FGDC geospatial metadata standard is to significantly enhance the capability of the FGDC metadata standard to be used to create effective, meaningful descriptions of biological data and information. This means that more existing biological data will be documented according to FGDC standards, and more individuals and institutions interested in sharing, accessing, and using biological data can participate in the NBII/ NSDI federation.